ADVANCED ROBOTICS AND AUTOMATION: KEY CONSIDERATIONS FOR HUMAN INTERACTION AND TRUST

With growing autonomy in advanced robotic systems, the language used to describe the relation between a human and robot might shift from ‘use’ to ‘interaction’. The quality of this interaction can impact a number of occupational safety and health (OSH) related factors and therefore should be a point of consideration when designing these systems. Within scientific literature, numerous robotic interaction design aspects, including cognitive ergonomics, are discussed in relation to different OSH aspects. They can, for example, be related to the outward appearance and embodiment of the robotic system, robot behaviour and movement or interaction as well as communication styles and channels. These different interaction design aspects are to varying degrees associated with OSH risks and opportunities. The similarity all interaction design research shares is the attempt to identify attributes and characteristics that enable a smooth and natural interaction. Furthermore, the overall aim is to increase the feeling of wellbeing, acceptance, trust, positive emotions, a positive user experience or workflow, while avoiding negative outcomes such as stress[1]. However, robotic design aspects are not stand-alone considerations but must always take into account the addressed context and working task.

Anthropomorphic robot design

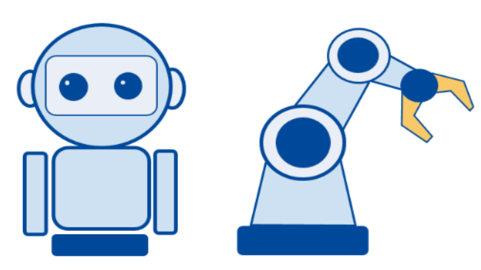


Figure 1: Robots with prevalent anthropomorphic features (eyes, ears) and without

Fictional robotic systems often display significant anatomical human likeness. While in reality this represents only a small fraction of how robotic systems are designed. The aspect of embodiment and more precisely anthropomorphic robotic design is addressed with great interest in scientific literature. Anthropomorphic design features like eyes or facial expressions can foster a more natural interaction, acceptance and likability especially in social robotics[2]. However, anthropomorphic design is not only limited to embodied features like facial components or body structure. It can also relate to robotic movements or communication strategies. While greater human likeness can have positive effects on trust towards robots, using human social cues as a means for a smooth human-robot interaction does not only have advantages. To begin with, there is no linear relationship between robotic anthropomorphism and associated likability or acceptance. Once a robot has reached a certain degree of human likeness, it can rather cause strong negative emotions like eeriness, often explained by the uncanny valley[3]. In addition, but probably more important are the negative consequences anthropomorphic design can impose on human expectations but also on task performance. Some features trigger **human expectations** regarding robotic capabilities and behaviour[4]. If a system has eyes, operators might expect the robot to perceive visual cues; if a robot has ears, auditory perception might be assumed. In general, anthropomorphic features may lead workers to subconsciously attribute human qualities like reasoning to the robot[5], and these assumptions can influence their interaction with the system.

> ***An operator’s expectation of robotic capabilities can be influenced by anthropomorphic design choices.***

---

[1] Honig, S. S., & Oron-Gilad, T. (2018). Understanding and resolving failures in human-robot interaction: Literature review and model development. *Frontiers in Psychology*, *9*, 861. https://doi.org/10.3389/fpsyg.2018.00861

[2] Fink, J. (2012, October). Anthropomorphism and human likeness in the design of robots and human-robot interaction. *International Conference on Social Robotics* (pp. 199-208). Springer. https://doi.org/10.1007/978-3-642-34103-8_20

[3] Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98-100. https://doi.org/10.1109/MRA.2012.2192811

[4] Złotowski, J., Proudfoot, D., Yogeeswaran, K., & Bartneck, C. (2015). Anthropomorphism: opportunities and challenges in human–robot interaction. *International Journal of Social Robotics*, *7*(3), 347-360. https://doi.org/10.1007/s12369-014-0267-6

[5] Murashov, V., Hearl, F., & Howard, J. (2016). Working safely with robot workers: Recommendations for the new workplace. *Journal of Occupational and Environmental Hygiene*, *13*(3), D61-D71. https://doi.org/10.1080/15459624.2015.111670

This can result in irritation or even a perceived lower reliability in industrial settings[6] if the technology cannot deliver on the prescribed qualities. Hence, when designing a robotic system, any anthropomorphic cue should always be considered with possible attributed functions. If those anthropomorphic cues do not serve a task related purpose, that is improve task performance, or support the coordination of the interacting partners, they should generally not be applied. Especially industrial robotics need to take this into account. In this setting, anthropomorphism can improve the workflow; for example, during physical pick and place tasks, anthropomorphic robot movements compared to purely robotic movements seem to enable humans to respond to the movement significantly faster and with greater accuracy[7]. However, in industrial settings, a misattribution of robotic capabilities by the human, especially for physically dangerous tasks, could pose a risk to workers. For example, incorrectly interpreted robotic movement intentions could possibly lead to collision. In social settings, the misattribution of capabilities based on anthropomorphism might hinder or halt the interaction all together. The interaction with a social robot that presents both eyes and ears, but is only capable of computing written text based input, can be perceived as irritating or frustrating.

It is challenging to identify the characteristic features distinguishing advanced robotics from non-robotic automation technologies to derive robotic specific OSH risks and challenges. While some aspects, to some extent, do also apply to non-robotic technologies, the dimension of interaction design, and more precisely the aspect of anthropomorphic robot design, seems distinct to the technology of advanced robotics. Anthropomorphic cues can benefit the interaction process between humans and robots, especially in social robotics. Especially in relation to physical tasks, anthropomorphic features pose the risk of irritation and false attributions if they are not explicitly dedicated to a task-relevant function.

Interaction principles and transparency in HRI

Transparency in HRI is important. However, too much information might cause an information overload and hinder processing of critical information.

The anthropometric design of advanced robotics embodiment can be regarded as quite robot specific, even unique. However, the application of well-known, more general design principles will also benefit the overall interaction process in the case of robotic systems. In relation to interface and interaction design, more traditional aspects and the overall existing knowledge on advantages of ergonomic design apply. One standard to consult when addressing interaction design are the interaction principles (former dialogue principles) formulated in the EN ISO 9241-110. Interaction principles and general design recommendations can guide the development and evaluation of user interfaces, leading to improved usability. The priority with which the interaction principles are applied depends on the purpose of the system, the users of the system, the tasks, the environment, the specific interaction technique used and the consequences arising from use. The way they are then deployed depends on what kind of system is being used. The seven principles: 'Suitability for the user's tasks'; 'Self-descriptiveness'; 'Conformity with user expectations'; 'Learnability'; 'Controllability'; 'Use error robustness'; and 'User engagement' form a basis on which users can evaluate their interaction with a system. While they have been identified to be important and useful for designing system interaction in the context of 'Industry 4.0'[8] and have proven to be an adequate tool for user evaluation of robotic systems[9], the literary basis on their application on robotic system interaction is still rare. Especially the new degree of autonomy that AI-based systems and advanced robotics bring into a workplace introduce a new quality to the interaction, which could be assessed and improved by applying the interaction principles early in the development process. Furthermore, new work environments will challenge cognitive abilities such as coordination, supervision and decision-making more than previous physical

---

[6] Roesler, E., Onnasch, L., & Majer, J. I. (2020, December). The Effect of anthropomorphism and failure comprehensibility on human-robot trust. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 107-111. SAGE Publications. https://doi.org/10.1177/1071181320641028

[7] Kuz, S., Faber, M., Bützler, J., Mayer, M. P., & Schlick, C. M. (2014, July). Anthropomorphic design of human-robot interaction in assembly cells. *Advances in The Ergonomics in Manufacturing: Managing the Enterprise of the Future,* AHFE Conference (pp. 265-272).

[8] Fischer, H., Engler, M., & Sauer, S. (2017, July). A human-centered perspective on software quality: acceptance criteria for work 4.0. In A. Marcus & W. Wang (Eds.), *International Conference of Design, User Experience, and Usability: Theory, Methodology, and Management* (pp. 570-583). Springer, Cham. https://doi.org/10.1007/978-3-319-58634-2_42

[9] Rosen, P. H., Sommer, S., & Wischniewski, S. (2018, August). Evaluation of human-robot interaction quality: A toolkit for workplace design. *Proceedings of the 20th Congress of the International Ergonomics Association* (pp. 1649-1662). Springer, Cham. https://doi.org/10.1007/978-3-319-96071-5_169

tasks[10]. For this reason, cognitive and sensorial aid needs to be provided to prevent information overload and its negative effects on the operator, also when mainly performing physical tasks. The quality and effectiveness of the addressed cognitive and sensorial aids is directly related to the aspect of interaction design and how well a system incorporates the described design principles. Especially as robotic systems expand in capabilities and autonomy, developers and legislators need to consider the facet of responsibility and accountability in the interaction. Humans hold robots accountable for their mistakes[11], at least more than other objects. Related to robot autonomy is the degree of system transparency. Few studies have specifically examined the impact of transparency on human-robot interaction so far. Transparency had a greater influence on user perceptions of the robot when the robot had greater autonomy[12]. Users placed greater blame on the robot and less blame on others when errors occurred in the work process. This suggests that transparency increases in importance as a robot's autonomous capabilities increase. Transparency stands for a more comprehensive treatment of information, including system state indication that an operator may need when dealing with autonomous systems; especially under high stress, workload, or uncertainty. If transparency is lacking, the user may view the robotic system as unreliable when in reality the provided information is misunderstood or not presented in a sufficient form. In that case, even routine behaviours can be interpreted as errors if the operator lacks the information to understand the reasoning process behind the actions[12]. In addition, given the imperfect track record of automation, it is imperative that researchers and developers consider the **transparency** of human-robot interaction to allow individuals to properly assess their reliance on these systems, particularly as technology gets more complex and is used in increasingly complex scenarios[13]. However, one should not simply assume that more information delivered by the system is necessarily better for the user. Too much information might not increase the transparency of a system, but lead to an information overload and result in an inability to select and process critical information[14]. Hence, creating sufficient transparency in human robot interaction is an important yet complicated endeavour with noticeable consequences for the interaction between operators and the system.

Incorporating **design principles** to a sound state, providing sufficient system transparency, or even enabling individualised interaction strategies that take into account very personal and individual preferences and characteristics will support a seamless system interaction. Technical requirements for individualisation or a smooth and user-friendly interaction are often related to the deployment of a variety of robotic sensors. Being able to react and behave in the intended or tailored to an individual way, requires the robotic system to collect and analyse environmental data as well as data regarding the interacting human. However, seamless interaction supported by well-known design principles can conflict with adverse effects like infringing user's privacy or contributing to the feeling of alienation and loss of control[15]. These negative feelings, can occur when a robotic system adapts its behaviour autonomously without notifying the user, who might not able to predict and understand this adaption[15]. Again, this interaction design-related risk emphasises the importance of transparency regarding system operations, actions and behaviour to reduce potential adverse effects. Any sort of individual related data collection also risks the actual or perceived feeling of monitoring of the employees' performance and behaviour. This can lead to negative impacts on motivation, satisfaction, organisational trust or stress, even if principles of data protection according to the General Data Protection Regulation (GDPR)

---

[10] Rauch, E., Linder, C., & Dallasega, P. (2020). Anthropocentric perspective of production before and within Industry 4.0. *Computers & Industrial Engineering*, *139*. https://doi.org/10.1016/j.cie.2019.01.018

[11] Kahn Jr, P. H., Kanda, T., Ishiguro, H., Gill, B. T., Ruckert, J. H., Shen, S., Gary, H. E., Reichert, A. L., Freier, N. G., & Severson, R. L. (2012, March). Do people hold a humanoid robot morally accountable for the harm it causes?. *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 33-40). https://doi.org/10.1145/2157689.2157696

[12] Kim, T., & Hinds, P. (2006, September). Who should I blame? Effects of autonomy and transparency on attributions in human-robot interaction. *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 80-85). IEEE. https://doi.org/10.1109/ROMAN.2006.314398

[13] Lyons, J. B. (2013, March). Being transparent about transparency: A model for human-robot interaction. *AAAI Spring Symposium: Trust and Autonomous Systems*. https://www.aaai.org/ocs/index.php/SSS/SSS13/paper/viewFile/5712/6000

[14] Finomore, V., Satterfield, K., Sitz, A., Castle, C., Funke, G., Shaw, T., & Funke, M. (2012, September). Effects of the multi-modal communication tool on communication and change detection for command & control operators. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1461-1465). SAGE Publications. https://doi.org/10.77/1071181312561410

[15] Fronemann, N., Pollmann, K., & Loh, W. (2021). Should my robot know what's best for me? Human–robot interaction between user experience and ethical design. *AI & SOCIETY*, 1-17. https://doi.org/10.1007/s00146-021-01210-3

are met[16]. Experts point out that especially cooperative or collaborative interaction scenarios could benefit from the advances in autonomous adaption of robotic behaviour; reacting for example to decreasing strength levels during a continuous task. However, the possible risks must not be neglected.

Trust as a key aspect in HRI

Adequate trust in the robot is vital. Excessive trust can lead to neglect, while mistrust can deter workers from using the robot.

Trust in automation, regardless of the specific automation technology, automation level or particular task is an important factor in human-machine interaction and often determines automation usage[17]. There are a vast number of definitions of trust that also stem from different disciplines, each emphasising different aspects, describing trust as a belief, attitude, intention or behaviour. Trust can be understood as 'the attitude that an agent [i.e. advanced robotics] will help achieve an individual's goal in a situation characterised by uncertainty and vulnerability'[18]. An adequate level of human trust towards the interacting system promotes an appropriate system use[19]. Extreme forms of trust can lead to adverse effects. Over-releance or excessive trust, for example, can lead to automation complacency[20]. However, insufficient trust may lead to neglect of the technology[18]. Overtrust and distrust are always considered in relation to the actual system capabilities, also considered as calibrated trust[16], which is highly related to the reliability of an automation technology or robotic system[18]. In case trust is miscalibrated, problematic interactions occur. Humans are 'found to misuse (over- or under-rely on the robot), disuse (stop using the robot all together), or abuse (use the robot for purposes other than as designed) their robotic counterpart, respectively'[18]. Miscalibrations can have severe effects, for example, when an operator decides to no longer monitor a robotic system although some oversight is needed. In contrast, an operator might strictly monitor a robotic system, neglecting other relevant tasks.

Not enough trust in a robotic system can have negative consequences for the interaction; likewise, excessive trust in the robot can also have adverse consequences. If there is excessive trust, the duty of care towards the robot, for example, is neglected[21] and defects go unnoticed, leading to damage of the work piece or injuries to people. If the degree of trust that is placed in the robot matches the capabilities of the robot, efficient and safe collaboration can take place[20]. If operators trust the robot, they follow suggestions it makes and accept information provided[21], which means that informed decisions can be made. With a good trust fit, it is possible to benefit from the advantages of human-robot collaboration[22]. While the concept of appropriate trust in a robotic system seems intuitive, several researchers suggest that it is a multifactorial, highly individual concept that needs further research to be fully understood[19, 21, 23, 24, 25]. While these papers highlight the importance of adequate trust in a robotic system and the dangers of automation bias, concrete strategies on how to mitigate or avoid it are not yet developed. Explicit training in automation bias and adequate training with the systems are encouraged, but not further specified[24, 25].

---

[16] Funk, M., Rosen, P. H., & Wischniewski, S. Human-centered HRI Design–the More Individual the Better?. *Behavioural patterns and interaction modelling for personalized Human-Robot interaction*, 2020. https://doi.org/10.1145/3371382.3374846

[17] Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230–253. https://doi.org/10.1518/001872097778543886

[18] Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human factors*, *46*(1), 50-80. https://doi.org/10.1518%2Fhfes.46.1.50_30392

[19] Hancock, P. A., Kessler, T. T., Kaplan, A. D., Brill, J. C., & Szalma, J. L. (2020). Evolving trust in robots: Specification through sequential and comparative meta-analyses. *Human Factors*, *63*(7), 1196-1229. https://doi.org/10.1177/0018720820922080

[20] Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381-410. https://doi.org/10.1177/0018720810376055

[21] Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*(5), 517-527. https://doi.org/10.1177/0018720811417254

[22] Sanders, T., Oleson, K. E., Billings, D. R., Chen, J. Y., & Hancock, P. A. (2011, September). A model of human-robot trust: Theoretical model development. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*(1). 1432-1436. SAGE Publications. https://doi.org/10.1177/1071181311551298

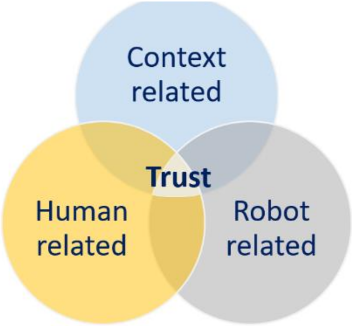
[23] Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors*, *58*(3), 377-400. https://doi.org/10.1177/0018720816634228

[24] Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121-127. https://doi.org/10.1136/amiajnl-2011-000089

[25] Papadimitriou, E., Schneider, C., Tello, J. A., Damen, W., Vrouenraets, M. L., & Ten Broeke, A. (2020). Transport safety and human factors in the era of automation: What can transport modes learn from each other?. *Accident Analysis & Prevention*, *144*, 105656. https://doi.org/10.1016/j.aap.2020.105656

Influential factors on trust

Figure 2: Three influential antecedents of trust in HRI



Results from different meta-analyses have shown that trust and its antecedents play an important role for OSH. There is agreement that antecedents significantly influencing human trust towards robotic systems can be human-, robot- or context-related and therefore have to be considered carefully when using robotic systems for the automation of tasks[19, 21, 23]. Among human-related factors exerting effects on human trust towards robots are user satisfaction, expectancy and comfort with robots. Robot-related factors can further be divided into performance factors, like reliability and failure rates, and attributes, such as the degree of anthropomorphism and physical appearance. Regarding performance, the related factor of dependability shows a negative correlation with trust; this means that as dependability increases, trust decreases. Regarding robotic attributes, literature indicates that robot personality factors, such as positive facial expressions, empathy, likability and sociability, show positive relations with trust[19]. As for context-related factors, the aspect of team collaboration, describing the constitution of a team, indicates a positive relation to trust. Furthermore, literature suggests that task difficulty has a significant effect on trust. Tasks that are more difficult, evoke higher levels of trust towards robotic systems. A reason for this might be reduced workload for the human as the robot might take over some challenging tasks. The reliability of a robot also has a positive and significant impact on trust. The more reliable a robot is, the more the human can rely on it to perform in a way that meets ones expectations[20]. Another significant source of influence revealed by literature is **proximity**. The closer the location of the robot to the human, the higher the degree of trust (Hancock et al., 2020). This finding is especially relevant for remote, teleoperating scenarios, which then especially should consider other trust enhancing aspects. Furthermore, the aspects of experience with a robot and anthropomorphism show a positive relationship with trust, indicating that more experience and higher degrees of anthropomorphism lead to greater trust towards the robotic system. Human trust towards the robot is also greater when robots fulfil user expectations and when users experience greater satisfaction. Overall, Hancock and colleagues conclude that factors relating to the robot have the strongest impact on trust compared to factors relating to the human. Within the robot-related antecedents, a robot’s attributes and its performance have the strongest impact on trust. However, the adverse effects of excessive levels of trust must not be neglected, especially when combined with a limited understanding of how the automating technology operates. This can lead to dangerous situations like unexpected behaviours, not recognizing automation failure or too slow responses to automation failure[25]. This aspect is also closely linked to the issue of adequate training in relation to the use of automation technology.

Conclusion and recommendations

As robotic systems increase in autonomy and capability, we also observe a linguistic shift from saying that a worker ‘uses’ a technology to a worker ‘interacts’ with it. This interaction is influenced by a variety of factors, which may be related to the work context, such as proximity, related to the robotic system in form of its appearance or behaviour, or related to the humans themselves and the trust they have in the robot.

As robotic systems become more advanced, more system characteristics pose OSH related opportunities and risks likewise. **Anthropomorphic design** or **individualisation features**, for example, can **ease and improve** the **interaction**. However, negative consequences for OSH, like **false expectations**, **surveillance** or **data privacy issues** can arise. A main attempt should therefore be made to carefully **counterbalance** risks and opportunities associated with specific design elements.

Even though advanced robotics are a comparatively new technology, we recommend involving **proven design principles**, like the interaction principles aligned with Human System Interaction Ergonomics Standards such as EN ISO 9241-110, when creating working systems, as they have a universal value in creating a human-centred system. Efficient and precise communication about the **robots capabilities** is also vital for successful interaction. This also implies **expectation management** when designing a robotic system. Including design features, which are not linked to a specific function, can pose a risk for workers. Specifically anthropomorphic design can lead to false expectations about the robots capabilities; hence, design elements like eyes or ears should only be included when the robot has a way to process the type of information associated with this input channel. **Transparency of a**

system's capabilities is therefore a key factor to facilitate human-robot interaction successfully. However, there needs to be a balance between creating sufficient transparency and causing information overload. What needs to be communicated to the worker regardless is if and what sort of **individual data** is collected by the robotic system. The way an interaction is designed can have major implications for the workers mental and physical wellbeing. Hence, it is important to create a system that provides transparency to the user and fosters **trust**, without overwhelming the worker.

Authors: Patricia Helen Rosen, Federal Institute for Occupational Safety and Health (BAuA), Eva Heinold, Federal Institute for Occupational Safety and Health (BAuA), Elena Fries-Tersch, Milieu Consulting SRL, Dr Sascha Wischniewski, Federal Institute for Occupational Safety and Health (BAuA).

Project management: Ioannis Anyfantis, Annick Starren, Emmanuelle Brun (EU-OSHA).

This policy brief was commissioned by the European Agency for Safety and Health at Work (EU-OSHA).Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect the views of EU-OSHA.